Performance of 3D-RISM-KH in predicting hydration free energy: effect of solute parameters

Roy, Dipankar; Kovalenko, Andriy

Publisher's version / Version de l'éditeur:
https://doi.org/10.1021/acs.jpca.9b01623
The Journal of Physical Chemistry A, 123, 18, pp. 4087-4093, 2019-04-17

THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

Cite This: J. Phys. Chem. A 2019, 123, 4087−4093

pubs.acs.org/JPCA

Performance of 3D-RISM-KH in Predicting Hydration Free Energy: Effect of Solute Parameters

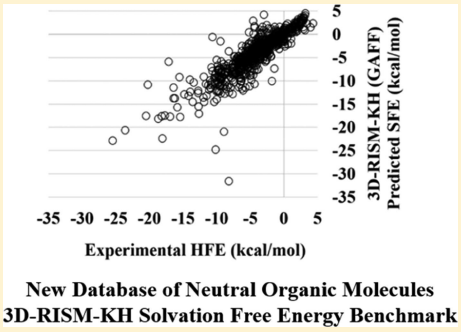
Dipankar Roy[†] and Andriy Kovalenko*[,†,‡]

[†]Department of Mechanical Engineering, University of Alberta, 10-203 Donadeo Innovation Centre for Engineering, 9211-116 Street NW, Edmonton, Alberta T6G 1H9, Canada

[‡]Nanotechnology Research Centre, 11421 Saskatchewan Drive, Edmonton, Alberta T6G 2M9, Canada

Supporting Information

ABSTRACT: The three-dimensional reference interaction site model molecular solvation theory with the Kovalenko−Hirata closure relation has been shown to produce excellent solvation characteristics for a large class of (bio)chemical systems in solution. Correct calculation of hydration free energy is central to successful application of any solvation model. In order to find out the best possible force-field parameters to be used for hydration free energy calculation with the aforementioned theory, we have developed an extended database containing a large number of experimental solvation free energies available in the current literature and used a plethora of theoretical models for assessment. The general Amber force field was found to perform satisfactorily, whereas special care should be taken in solute charge assignment with the universal force field.



INTRODUCTION

The knowledge of hydration free energy (HFE) is essential for applications spanning the biological, chemical, and materials science fields. While the experimental determination of molecular solvation energy may seem straightforward, it often gets complicated for practical applications. Predictive methods that can help produce reliable solvation energy are in demand, as they eliminate the need for a large number of experiments.[1,2] The HFE, in fact solvation energy in all solvents, is tightly coupled with the electronic properties of the molecule of interest.[3] Thus, a logical point is to use quantum chemical methods to calculate HFE. While the implicit solvation models based on the integral equation formalism are impressive for wide-spread applications,[4,5] it is impractically time consuming to screen a large molecular database, for example, aimed for drug development activities. Further, results obtained by the continuum models can be improved by explicitly considering the first solvation shell in a solvation thermodynamics calculation.[6,7] Such a cluster continuum model is well known in chemistry.[8] While this methodology is a very productive option, the consequential large supermolecular structure often becomes too big for calculations with a very high precision. Thus, applications of this model so far have been limited to chemical reactions and thermochemical property calculations for molecules with reasonable sizes.

Over the years, several other solvation models were developed to predict/calculate hydration/solvation free energy with variable degrees of precision, which is often governed by the nature of intended application(s).[9] For example, semiempirical HFE-based methods were often used for applications in physicochemical property evaluations.[10−12] The physics-based methods were, on the other hand, designed to develop understanding of a solvation property based on the molecular property of the liquid solvent. For all solvent model development purposes, an essential component is a robust training set. A chemically diverse data set yields a more robust solvation model. An excellent example of this fact is the Minnesota solvation models.[13] Significant efforts were devoted to develop well-balanced solvation free energy databases containing carefully curated experimental data points.[14−16]

The three-dimensional reference interaction site model (3D-RISM) is a first-principle statistical thermodynamics-based method which uses moderate computational resources to calculate thermodynamics associated with solvation in molecular liquids. The molecular solvation theory with the integral equation formalism is based on the modified Ornstein−Zernike theory for dimensional reduction of molecular liquids.[17] In the RISM theory, a molecular interaction potential force field is used to analyze thermodynamic properties based on correlation functions between molecules in solution from the ensuing thermodynamic ensemble. The essential part of the 3D-RISM formalism produces the 3D distribution functions ($g_\gamma(\mathbf{r})$) of solvent interaction sites around a solute molecule of arbitrary shape, which is correlated to the total correlation function ($h_\gamma(\mathbf{r})$) at position $\mathbf{r}$ around the solute molecule as

$$h_\gamma(\mathbf{r}) = \sum_\alpha \int \mathrm{d}r' c_\alpha(\mathbf{r} - \mathbf{r}')\chi_{\alpha\gamma}(r')$$

where $h_\gamma(\mathbf{r})$ is the 3D total correlation function of solvent site γ related to $g_\gamma(\mathbf{r})$ by $g_\gamma(\mathbf{r}) = h_\gamma(\mathbf{r}) + 1$, and $c_\gamma(\mathbf{r})$ is the 3D direct

Received: February 19, 2019
Revised: April 15, 2019
Published: April 17, 2019

DOI: 10.1021/acs.jpca.9b01623
J. Phys. Chem. A 2019, 123, 4087−4093

correlation function which has the asymptotics of the solute−solvent site interaction potential, $c_\gamma(\mathbf{r}) \approx -u_\gamma(\mathbf{r})/(k_B T)$; the site−site susceptibility of pure solvent $\chi_{\alpha\gamma}(r)$ is an input to the 3D-RISM theory; and indices α and γ enumerate all sites on all sorts of solvent species at temperature T and Boltzmann constant k_B.[18−21] A detailed description of the molecular solvation theory and derivations of the RISM theory are described elsewhere.[22] It is imperative to note that it is impossible to solve the 3D-RISM integral equation for viable solutions without the so-called closure relations. A closure relation connects the 3D total and direct correlation functions, $h_\gamma(\mathbf{r})$ and $c_\gamma(\mathbf{r})$, via a 3D solute−solvent interaction potential. The most successful closure relation was developed by Kovalenko and Hirata (KH) which keeps the right asymptotic behavior while taking care of spatial distributions in the regions with solvent density enrichment ($g_\gamma(\mathbf{r}) > 1$) as well as regions of solvent density depletion ($g_\gamma(\mathbf{r}) < 1$) automatically.[23,24] It is important to note that the 3D-RISM-KH theory underestimates the height of strongly associative peaks of the 3D distribution function while at the same time widening the peaks; hence, the resultant thermodynamic properties calculated by this theory come out quite accurate.[25]

A major weakness of the 3D-RISM-KH theory is the correct calculation of the HFE.[26] Several modifications and/or enhancements of the 3D-RISM theory were reported to overcome this issue. The most popular correction scheme is based on the so-called "universal correction" (UC) based on the regression model built on the Gaussian fluctuation (GF) functional and partial molar volume (PMV) obtained from a 3D-RISM calculation.[27] The correction is of the form: $\Delta G_{UC} = \Delta G_{GF} + a \times \text{PMV} + b$; where a and b are determined from regression analysis. The units for these regression coefficients are kcal/mol/Å^3 and kcal/mol, respectively.

This correction yielded excellent correlation between experimental and predicted HFE.[28−31] Another correction explores the shortcomings of the cavitation free energy calculations in the scheme and accounts for them empirically.[32] Correction schemes based on molecular dynamics simulation were also reported.[33] The solvation free energy pressure correction is another way to account for an error in internal pressure in a 3D-RISM calculation.[34] A closer look into these correction schemes yields several interesting facts. First, the UC correction scheme is dependent on the size and accuracy of the database used for calibration, and the bigger the database, the better the suitability of correlation factors. Second, partial charges on the solute molecule are important, and hence, suitability of atomic population partitioning schemes need to be standardized for applications in HFE calculation using the 3D-RISM-KH method.

In this work, we have developed a larger database containing experimentally determined HFEs of small neutral organic molecules to examine the performance of the 3D-RISM-KH theory in training and predicting HFE. We have tested different atomic charge generation schemes to compare their effect in predicting HFEs.

DATABASE PREPARATION AND COMPUTATIONAL METHODS

The data set containing experimental HFEs of small organic molecules was collected from the previously reported FreeSolv database of Mobley and co-workers, the database reported by Marenich et al. for calibrating the performance of SMx solvation models, and the HFE database from Guthrie.[15,16,35] All molecules that appeared multiple times were removed, resulting in a database of 822 molecules. For multiple occurrences of a molecule, data from the FreeSolv database are used. Additional HFE's of drug molecules were collected from the literature.[36−39] See the Supporting Information for details of molecules with the experimental HFE and functional groups distribution used in this study. 75% of compounds were assigned randomly for the training set. The remaining molecules were used to build the test set. The details of these sets are provided in the Supporting Information. The UC correction schemes were obtained by ordinary linear regression analysis. For this purpose, the functional form optimized is $\Delta G_{GF} - \Delta G_{Exp} = a \times \text{PMV} + b$; the correlation coefficients (a and b) are reported in the Supporting Information. The performance of the models was estimated by mean absolute deviation (MAD) and relative mean square error (RMSE). The MAD (in kcal/mol) is calculated as $\frac{\sum |a_e - a_0|}{n}$, where a_e and a_0 are experimental and predicted HFEs, respectively, and n is the number of data points. The RMSE (in kcal/mol) is calculated as $\sqrt{\frac{\sum (a_e - a_0)^2}{n}}$.

All the molecules were subjected to geometry optimization at the density functional theory (DFT) M06-2X/Def2-TZVPP level using the Gaussian 16 software package.[40−42] Optimized structures were verified as minima at the respective potential energy surface by vibrational analysis. These optimized geometries were used for the 3D-RISM-KH calculations (vide infra). The effect of water solvent continuum was incorporated via the SMD and conductor-like polarizable continuum (CPC) model solvation models as implemented in the Gaussian 16 software package.[35,43] The default setting of the SMD model in the Gaussian 16 package uses scaled Bondi radii, while the CPCM model uses the united atom model with universal force field (UFF)-based atomic radii to model the solute cavity. In addition to the Mulliken atomic population scheme, several atomic charge schemes were employed to generate partial atomic charges of the solute molecules. Partial atomic charges based on the natural population, CHelpG charges, Merz-Singh-Kollman (MK), and charge model 5 (CM5) were calculated using Gaussian 16 software at the abovementioned level.[44−47] The DDEC6 charges were calculated using the program Chargemol of Manz and co-workers using a wave function file generated at the M06-2X/Def2-TZVPP level of theory.[48−50] The choice of the Ahlrichs triple-ζ basis set is based on the fact that the present data set contains bromine and iodine atoms containing molecules. The Amsterdam density functional (ADF)[51] code provides an extended version of the 3D-RISM-KH code allowing a DFT grid to be used for the 3D-RISM calculations. However, a clear choice of a method has not yet been developed for this purpose. Hence, we have additionally tested the implementation of the 3D-RISM methodology in the ADF code using the local density approximation (LDA),[52] the hybrid B3LYP functional,[53−55] as well as the metahybrid M06-2X functional with both the double- and triple-ζ quality basis set. The SPCE water model[56] as provided in the ADF suite was used for solvent 3D correlation function calculations.

All the RISM calculations (1D- and 3D-) were performed in our locally modified code, a working version of which is made available in the AmberTools program.[57] The 3D-RISM-KH calculations were performed using a homogenous cubic grid of 128 × 128 × 128 points in the box of size 64 × 64 × 64 Å^3 to account for the solute with a few solvation layers. The convergence accuracy of the modified direct inversion in the

DOI: 10.1021/acs.jpca.9b01623
J. Phys. Chem. A 2019, 123, 4087−4093

iterative subspace solver was set to 10^{-4}. The site−site susceptibility function of bulk water solvent is obtained from the dielectrically consistent RISM theory (DRISM) theory with the KH closure. For the 1D-RISM calculations with bulk water solvent, we have used the SPCE, TIP3P, TIP4P, and TIP5P point charge-based models of water. Another modification of the SPCE model as implemented in the ADF package, denoted as SPCE2, was tested, too. Details of these solvation models as applied to the 1D-RISM applications are also available from the AMBER molecular dynamics code. The solute molecules were prepared using the tLEAP program from the Antechamber software and were manipulated to prepare suitable input decks for 3D-RISM-KH calculations using simple computer scripts. This process assigns general Amber force field (GAFF) parameters for the solute molecules.[58,59] Additionally, we have tested the UFF parameters for the solute[60] with different charge schemes for suitability in the HFE calculations.

The cheminformatics-based point charges were often used for rapid calculation of molecular descriptors, and hence, for property/activity prediction. A number of them were implemented in the cheminformatics software package OpenBabel.[61] We have tested (i) atoms in molecule charges, (ii) natural populations, (iii) EQEq, (iv) QEq, and (v) QTPIE atomic charges for all the solutes generated via the OpenBabel package.[62−66] A pictorial description of the workflow adopted in this report is presented in Figure S1 in the Supporting Information.

■ RESULTS AND DISCUSSION

In this work, we strive to provide the best possible solution for calculating solvation free energy using the 3D-RISM-KH theory in terms of force-field parameters and atomic charges combinations. Our objective stems from the fact that a successful 3D-RISM calculation of HFE requires assignment of force-field parameters (σ and ε) and atomic charges for solute molecules. Thus, the error in calculations will depend on these components. While assigning GAFF parameters for small molecules is widely practiced, this force field is not suitable for metal containing systems. The UFF, on the other hand, is designed for the entire periodic table but lacks proper charge assignments. Hence, we have decided to analyze performance of this force field with different partial atomic charge generation schemes for solutes in predicting HFEs.

To avoid large tables with individual-computed HFEs in this section, all the predicted HFEs were reported in the Supporting Information. The present discussion is divided into three sections. First, performance of different water force field parameters on the 3D-RISM-KH prediction of HFE is discussed. The second section deals with the choice of the atomic charge generation schemes on the accuracy of solvation free energy prediction. The final section deals with an overall comparison among solvation energy predicted via the 3D-RISM-KH theory and continuum solvation models (SMD, CPCM) with reference to the experimental HFE.

The performance of different water force fields (Table 1) in predicting HFEs is found to be dependent on the choice of the parameters for solutes. For instance, the SFEs predicted with the TIP5P water model for DRISM calculation of water yielded the performance marginally better than the SPCs, TIP3P, and TIP4P models for solutes parametrized with the GAFF and UFF parameters, except for the UFF parameters with DDEC6 charges for the solute. We have noticed comparably longer calculation times for convergence with the TIP4P and TIP5P solvent parameters in comparison to the other three models. Figure 1 attests to these findings. The performance of the GAFF is comparable amongst different water force fields used. The prediction of HFEs for the molecules containing S and/or P atoms in a single molecule with the GAFF parameters for solutes is below par across all water force fields, whereas the solutes with the UFF parameters and DDEC6 charges worked uniformly across all the classes of molecules. The overall performance of the GAFF is the best for all of the water models used, as shown by the MAD (in kcal/mol) in the predicted HFEs. The UFF for solutes worked well, too, with the AM1-BCC charges with marginally larger MADs.

Table 1. Lennard-Jones Parameters of the Water Models Used in This Report

parameters	SPCE	SPCE2	TIP3P	TIP4P	TIP5P
σ_O/Å	3.1658	3.166	3.1507	3.1536	3.1199
ε_O/kcal/mol	0.1553	0.1544	0.1520	0.1550	1.6000
σ_H/Å	0.4	0.7	0.4	1.2392	1.2055
ε_H/kcal/mol	0.046	0.046	0.046	0.0155	0.0160

The choice of atomic charges is paramount for application of the UFF parameters for solutes in a 3D-RISM calculation. We have tested several charge schemes to find out the best candidate. Our findings are summarized in Figure 2. First, the cheminformatics-based charges showed abysmal performance with the UFF for both the training and test sets. A marginally better performance is observed for the QTPIE charges for solute molecules and the TIP5P water model (MAD of 3.0 kcal/mol). The AM1-BCC charge model worked much as the various DFT-calculated partial charge generation methods (Figure 2). The DDEC6 level charges did not improve the prediction of solvation energy. It is important to note that in the present calculation setup, the generation of DDEC6 charges was carried out by two layers of calculations, namely, generation of the molecular wave function file via electronic structure calculations and postprocessing the wave function file to generate partial charges; hence, significant computing time is added to the entire process of the 3D-RISM calculations. The atomic charges calculated by charges fitted to the electrostatic potential according to the MK scheme and CHelpG scheme have performed satisfactorily. These two population schemes were used with a constraint to reproduce the dipole moment. Interesting to note is that the charge model by Truhlar−Cramer (CM5) performed uniformly well across all the force-field models of water (MAD of 1.8 kcal/mol).

The implementation of the 3D-RISM-KH code in the ADF package requires the σ and ε parameters for each atom of the solute along with atomic coordinates. In our calculations with the ADF package, we have used the TZP basis set. The choice of the density functional did not show much influence in predicting HFEs. For instance, the MADs for the calculations with the LDA, B3LYP, and M06-2X functionals are in the range of ∼3.7 kcal/mol. We would like to point that the use of a double-ζ basis set for these calculations yielded negative PMV for several systems, which is unphysical and points to erroneous calculations of the distribution functions (Figure 3). The PMV is an important thermodynamic property related to a solute orientation in solution as well as to the solute−solvent interactions. In the Kirkwood−Buff theory which is extended further to the 3D-RISM-KH theory, the PMV is expressed as a function of isothermal compressibility of pure solvent (positive for pure water) and the solute−solvent radial distribution

DOI: 10.1021/acs.jpca.9b01623
J. Phys. Chem. A 2019, 123, 4087−4093

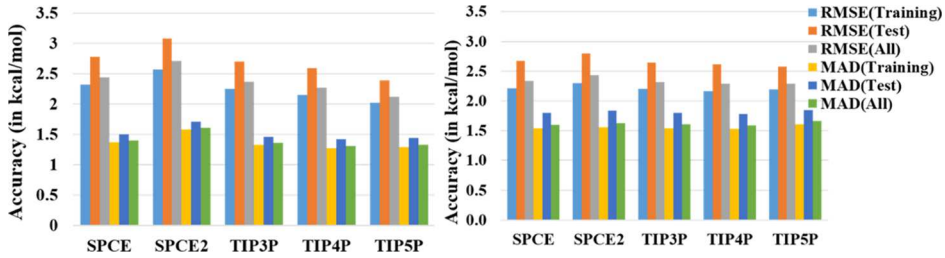


Figure 1. Performance of the GAFF for solute molecules used in predicting HFE in combination with five different force field parameters for water solvent by the RISM-KH theory.

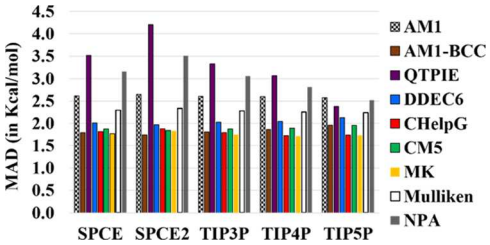


Figure 2. Performance of the UFF with different partial atomic charge assignments for solute molecules used in predicting HFE in combination with five different force field parameters for water solvent by RISM-KH theories.

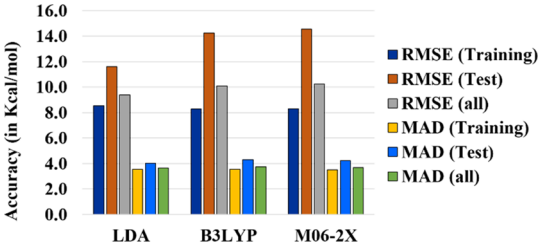


Figure 3. Performance of the UFF for solute molecules used in predicting HFE in combination with the SPCE2 force-field parameters for water solvent by the 3D-RISM-KH theory as implemented in the ADF software package.

function.[67] A negative PMV can be obtained for solutes with strong ionic charges (e.g., Li^+) and zwitterions.

Finally, a direct comparison of the performances of the continuum solvation models, CPCM and SMD, with the 3D-RISM-KH points to several findings (Figure 4). The performance of the SMD solvation model is the best for the entire data set. This is attributed to the fact that a majority of the molecules reported here were a part of the database used for parameterizing the SMD model. The performance of the CPCM model (MAD of 1.8 kcal/mol) is higher than the best prediction of the 3D-RISM-KH theory in combination with the GAFF parameters for

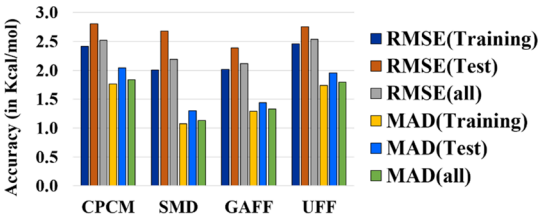


Figure 4. Comparative performance of the 3D-RISM-KH theory in predicting HFE in combination with the GAFF and UFF with the continuum solvent models CPCM and SMD.

solutes and TIP5P for the water model (MAD of 1.3 kcal/mol) and is comparable to that predicted by the UFF parameters for solutes with the AM1-BCC charges (MAD ∼1.8 kcal/mol).

The correlations of the 3D-RISM-KH (GAFF)-predicted HFEs with the experimental results are excellent. The same for the UFF-predicted solvation free energy did not perform, whereas the GAFF did (Figure 5). Further studies are in progress to extend the findings for systems containing transition metal(s). The predicted HFEs for FreeSolv database is similar to that of the training set, probably owing to the comparable sizes. A direct comparison of the previously reported performances[27,32] of the 3D-RISM-KH theory in predicting HFEs further points to the adequacy of the methodology that we present. For instance, our calculations on a data set consisting 504 small molecules reported by Truchon et al. yielded an MAD of 0.89 kcal/mol for calculations with the GAFF for solutes and the TIP3P water model (the previously reported mean unsigned error by the authors was in the range of 0.91−0.95 kcal/mol using various UC correction schemes).[32] Another data set, albeit small, reported by Palmer et al.[27] yielded an MAD of 0.71 kcal/mol for all the 185 molecules with the GAFF parameters for solutes and the SPCE water model, which is comparable to the previous report. For all the calculations, we have employed the lowest energy conformations of the solute molecules as obtained from the conformation analysis and subsequent energy calculation at the AM1 level with the MOPAC semiempirical quantum calculation package.[68,69] We have noticed large deviations in the predicted solvation energies (in the range of ±5 kcal/mol) from the benchmark data: for 28 compounds with the CPCM level, 19 compounds with the SMD level, 18 compounds with the 3D-RISM-KH (GAFF for solute, TIP5P for water), and 40 for the 3D-RISM-KH (UFF with AM1-BCC charges for solute, TIP5P for water) level of calculations. Majority of these compounds belong to the new addition to the data set and are drug/drug-like substances. Possible reasons for large fluctuation in the calculated values are due to large errors in experimental results and/or lack of conformational flexibility implied in the calculations. We intend to investigate the issue of the conformation flexibility in HFE calculations with 3D-RISM-KH calculation via the generalized solvation force extrapolation techniques in near future.

In this report, we have presented a comparative analysis of the performances of the GAFF and UFF for solute parametrization to predict HFE using the 3D-RISM-KH molecular solvation theory. For the purpose of comparison, we have developed an extended solvation free energy database by collecting the experimental HFE reported in the existing literature. Our results point out to the superior performance of the 3D-RISM-KH theory in predicting HFEs when applied with the GAFF in comparison to the CPCM. This particular level of the 3D-RISM-KH calculations performed at par with the SMD continuum

DOI: 10.1021/acs.jpca.9b01623
J. Phys. Chem. A 2019, 123, 4087−4093

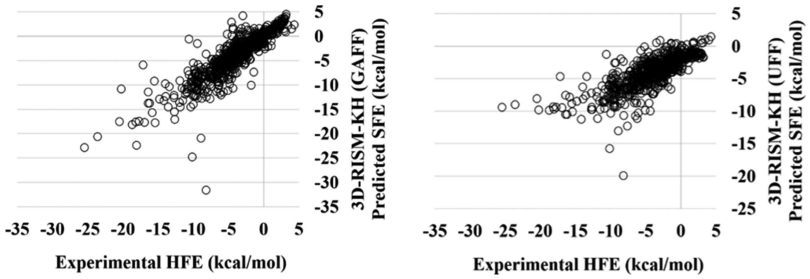


Figure 5. Comparative performances of the 3D-RISM-KH theory-based HFE prediction with the GAFF and UFF with reference to the experimental HFE database. The errors in calculating solvation free energy are provided in Table S2 in the Supporting Information.

solvation model, too. The 3D-RISM-KH calculations with the UFF parameters for the solute showed a slightly diminished performance with semiempirical charge models in HFE predictions but improved significantly upon coupling with the atomic charge generation schemes based on molecular electrostatic potentials (e.g., CHelpG, MK). The incorporation of the DDEC6 charge models did not improve the 3D-RISM-KH predictions in our hands but added significant computing time as additional molecular wave function files were to be processed to generate these charges. The key features of the 3D-RISM-KH calculations are the speed and accuracy, as shown in our present study.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.9b01623.

Experimental and computed HFE, identity of the systems, regression parameters, correlation plots between various prediction schemes, and complete citations of refs[42,57] (PDF)

Solvation energies in kcal/mol, experimental uncertainties in kcal/mol, and functional groups of various compounds (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: andriy.kovalenko@ualberta.ca. Phone: +1-780-940-7477.

ORCID

Dipankar Roy: 0000-0002-4703-0130

Andriy Kovalenko: 0000-0001-5033-4314

Author Contributions

The manuscript was written through equal contributions of all the authors. All the authors have given approval to the final version of the manuscript.

Funding

This work was financially supported by the NSERC Discovery Grant (RES0029477) and Alberta Prion Research Institute Explorations VII Research Grant (RES0039402).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge generous computing time provided by WestGrid (www.westgrid.ca) and Compute Canada/Calcul Canada (www.computecanada.ca). D.R. thanks Dr. Sergey Gusarov for useful discussions. The authors thank the Reviewers for valuable suggestions.

■ REFERENCES

(1) Hansen, N.; van Gunsteren, W. F. Practical Aspects of Free-Energy Calculations: A Review. *J. Chem. Theory Comput.* **2014**, *10*, 2632−2647.

(2) Bash, P.; Singh, U.; Langridge, R.; Kollman, P. Free Energy Calculations by Computer Simulation. *Science* **1987**, *236*, 564−568.

(3) Sato, H.; Sakaki, S. Comparison of Electronic Structure Theories for Solvated Molecules: RISM-SCF versus PCM. *J. Phys. Chem. A* **2004**, *108*, 1629−1634.

(4) Cramer, C. J.; Truhlar, D. G. Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chem. Rev.* **1999**, *99*, 2161−2200.

(5) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999−3094.

(6) da Silva, E. F.; Svendsen, H. F.; Merz, K. M. Explicitly Representing the Solvation Shell in Continuum Solvent Calculations. *J. Phys. Chem. A* **2009**, *113*, 6404−6409.

(7) Gupta, M.; da Silva, E. F.; Svendsen, H. F. Explicit Solvation Shell Model and Continuum Solvation Models for Solvation Energy and PKa Determination of Amino Acids. *J. Theor. Comput. Chem.* **2013**, *9*, 5021−5037.

(8) Pliego, J. R.; Riveros, J. M. The Cluster−Continuum Model for the Calculation of the Solvation Free Energy of Ionic Species. *J. Phys. Chem. A* **2001**, *105*, 7241−7247.

(9) Skyner, R. E.; McDonagh, J. L.; Groom, C. R.; van Mourik, T.; Mitchell, J. B. O. A Review of Methods for the Calculation of Solution Free Energies and the Modelling of Systems in Solution. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6174−6191.

(10) Anisimov, V. M.; Cavasotto, C. N. Hydration Free Energies Using Semiempirical Quantum Mechanical Hamiltonians and a Continuum Solvent Model with Multiple Atomic-Type Parameters. *J. Phys. Chem. B* **2011**, *115*, 7896−7905.

(11) Kříž, K.; Řezáč, J. Reparametrization of the COSMO Solvent Model for Semiempirical Methods PM6 and PM7. *J. Chem. Inf. Model.* **2019**, *59*, 229−235.

(12) McDonagh, J. L.; Nath, N.; De Ferrari, L.; van Mourik, T.; Mitchell, J. B. O. Uniting Cheminformatics and Chemical Theory to Predict the Intrinsic Aqueous Solubility of Crystalline Druglike Molecules. *J. Chem. Inf. Model.* **2014**, *54*, 844−856.

(13) Cramer, C. J.; Truhlar, D. G. A Universal Approach to Solvation Modeling. *Acc. Chem. Res.* **2008**, *41*, 760−768.

(14) Cramer, C. J.; Truhlar, D. G. SMx Continuum Models for Condensed Phases. In *Trends and Perspectives in Modern Computational Science*; Maroulis, G., Simos, T. E., Eds.; Lecture Series on Computer and Computational Sciences; Brill/VSP: Leiden, 2006; Vol. *6*, pp 112−139.

(15) Duarte Ramos Matos, G.; Kyu, D. Y.; Loeffler, H. H.; Chodera, J. D.; Shirts, M. R.; Mobley, D. L. Approaches for Calculating Solvation Free Energies and Enthalpies Demonstrated with an Update of the FreeSolv Database. *J. Chem. Eng. Data* **2017**, *62*, 1559−1569.

(16) Mobley, D. L.; Wymer, K. L.; Lim, N. M.; Guthrie, J. P. Blind prediction of solvation free energies from the SAMPL4 challenge. *J. Comput. Aided Mol. Des.* **2014**, *28*, 135.

DOI: 10.1021/acs.jpca.9b01623
J. Phys. Chem. A 2019, 123, 4087−4093

(17) Hansen, J.-P.; McDonald, I. R. Molecular Liquids. In *Theory of Simple Liquids*, 4th ed.; Hansen, J.-P., McDonald, I. R., Eds.; Academic Press: Oxford, 2013; pp 455−510.
(18) Chandler, D.; McCoy, J. D.; Singer, S. J. Density Functional Theory of Nonuniform Polyatomic Systems. I. General Formulation. *J. Chem. Phys.* **1986**, *85*, 5971−5976.
(19) Chandler, D.; McCoy, J. D.; Singer, S. J. Density Functional Theory of Nonuniform Polyatomic Systems. II. Rational Closures for Integral Equations. *J. Chem. Phys.* **1986**, *85*, 5977−5982.
(20) Beglov, D.; Roux, B. An Integral Equation To Describe the Solvation of Polar Molecules in Liquid Water. *J. Phys. Chem. B* **1997**, *101*, 7821−7826.
(21) Kovalenko, A.; Hirata, F. Self-consistent description of a metal-water interface by the Kohn-Sham density functional theory and the three-dimensional reference interaction site model. *J. Chem. Phys.* **1999**, *110*, 10095−10112.
(22) Kovalenko, A. Multiscale Modeling of Solvation. In *Springer Handbook of Electrochemical Energy*; Breitkopf, C., Swider-Lyons, K., Eds.; Springer Berlin Heidelberg: Berlin, 2017; pp 95−139.
(23) Kovalenko, A.; Gusarov, S. Multiscale methods framework: self-consistent coupling of molecular theory of solvation with quantum chemistry, molecular simulations, and dissipative particle dynamics. *Phys. Chem. Chem. Phys.* **2018**, *20*, 2947−2969.
(24) Kovalenko, A.; Hirata, F. A Molecular Theory of Liquid Interfaces. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1785−1793.
(25) Perkyns, J. S.; Lynch, G. C.; Howard, J. J.; Pettitt, B. M. Protein Solvation from Theory and Simulation: Exact Treatment of Coulomb Interactions in Three-Dimensional Theories. *J. Chem. Phys.* **2010**, *132*, 064106.
(26) Sergiievskyi, V. P.; Jeanmairet, G.; Levesque, M.; Borgis, D. Fast Computation of Solvation Free Energies with Molecular Density Functional Theory: Thermodynamic-Ensemble Partial Molar Volume Corrections. *J. Phys. Chem. Lett.* **2014**, *5*, 1935−1942.
(27) Palmer, D. S.; Frolov, A. I.; Ratkova, E. L.; Fedorov, M. V. Towards a Universal Method for Calculating Hydration Free Energies: A 3D Reference Interaction Site Model with Partial Molar Volume Correction. *J. Phys. Condens. Matter* **2010**, *22*, 492101.
(28) Ratkova, E. L.; Chuev, G. N.; Sergiievskyi, V. P.; Fedorov, M. V. An Accurate Prediction of Hydration Free Energies by Combination of Molecular Integral Equations Theory with Structural Descriptors. *J. Phys. Chem. B* **2010**, *114*, 12068−12079.
(29) Ratkova, E. L.; Palmer, D. S.; Fedorov, M. V. Solvation Thermodynamics of Organic Molecules by the Molecular Integral Equation Theory: Approaching Chemical Accuracy. *Chem. Rev.* **2015**, *115*, 6312−6356.
(30) Luchko, T.; Blinov, N.; Limon, G. C.; Joyce, K. P.; Kovalenko, A. SAMPL5: 3D-RISM Partition Coefficient Calculations with Partial Molar Volume Corrections and Solute Conformational Sampling. *J. Comput. Aided Mol. Des.* **2016**, *30*, 1115−1127.
(31) Johnson, J.; Case, D. A.; Yamazaki, T.; Gusarov, S.; Kovalenko, A.; Luchko, T. Small Molecule Hydration Energy and Entropy from 3D-RISM. *J. Phys. Condens. Matter* **2016**, *28*, 344002.
(32) Truchon, J.-F.; Pettitt, B. M.; Labute, P. A Cavity Corrected 3D-RISM Functional for Accurate Solvation Free Energies. *J. Chem. Theory Comput.* **2014**, *10*, 934−941.
(33) Maruyama, Y.; Yoshida, N.; Tadano, H.; Takahashi, D.; Sato, M.; Hirata, F. Massively Parallel Implementation of 3D-RISM Calculation with Volumetric 3D-FFT. *J. Comput. Chem.* **2014**, *35*, 1347−1355.
(34) Sergiievskyi, V.; Jeanmairet, G.; Levesque, M.; Borgis, D. Solvation Free-Energy Pressure Corrections in the Three Dimensional Reference Interaction Site Model. *J. Chem. Phys.* **2015**, *143*, 184116.
(35) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378−6396.
(36) Perlovich, G. L.; Kurkov, S. V.; Kinchin, A. N.; Bauer-Brandl, A. Thermodynamics of Solutions III: Comparison of the Solvation of (+)-Naproxen with Other NSAIDs. *Eur. J. Pharm. Biopharm.* **2004**, *57*, 411−420.
(37) Perlovich, G. L.; Kurkov, S. V.; Kinchin, A. N.; Bauer-Brandl, A. Thermodynamics of Solutions IV: Solvation of Ketoprofen in Comparison with Other NSAIDs. *J. Pharm. Sci.* **2003**, *92*, 2502−2511.
(38) Perlovich, G. L.; Kurkov, S. V.; Bauer-Brandl, A. Thermodynamics of solutions. *Eur. J. Pharm. Sci.* **2003**, *19*, 423−432.
(39) Perlovich, G. L.; Volkova, T. V.; Bauer-Brandl, A. Towards an Understanding of the Molecular Mechanism of Solvation of Drug Molecules: A Thermodynamic Approach by Crystal Lattice Energy, Sublimation, and Solubility Exemplified by Paracetamol, Acetanilide, and Phenacetin. *J. Pharm. Sci.* **2006**, *95*, 2158−2169.
(40) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(41) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297−3305.
(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. *Gaussian16*, Revision B.01; Gaussian Inc.: Wallingford, CT, 2016.
(43) York, D. M.; Karplus, M. A Smooth Solvation Potential Based on the Conductor-Like Screening Model. *J. Phys. Chem. A* **1999**, *103*, 11060−11079.
(44) Reed, A. E.; Weinstock, R. B.; Weinhold, F. Natural Population Analysis. *J. Chem. Phys.* **1985**, *83*, 735−746.
(45) Breneman, C. M.; Wiberg, K. B. Determining Atom-Centered Monopoles from Molecular Electrostatic Potentials. The Need for High Sampling Density in Formamide Conformational Analysis. *J. Comput. Chem.* **1990**, *11*, 361−373.
(46) Besler, B. H.; Merz, K. M., Jr.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* **1990**, *11*, 431−439.
(47) Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. Charge Model 5: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases. *J. Chem. Theory Comput.* **2012**, *8*, 527−541.
(48) Manz, T. A.; Sholl, D. S. Chemically Meaningful Atomic Charges That Reproduce the Electrostatic Potential in Periodic and Nonperiodic Materials. *J. Chem. Theory Comput.* **2010**, *6*, 2455−2468.
(49) Manz, T. A.; Limas, N. G. Introducing DDEC6 Atomic Population Analysis: Part 1. Charge Partitioning Theory and Methodology. *RSC Adv.* **2016**, *6*, 47771−47801.
(50) Limas, N. G.; Manz, T. A. Introducing DDEC6 Atomic Population Analysis: Part 4. Efficient Parallel Computation of Net Atomic Charges, Atomic Spin Moments, Bond Orders, and More. *RSC Adv.* **2018**, *8*, 2678−2707.
(51) ADF 2014, SCM; Theoretical Chemistry; Vrije Universiteit: Amsterdam, The Netherlands. http://www.scm.com (accessed Feb 18, 2019).
(52) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58*, 1200−1211.
(53) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(54) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785−789.
(55) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623−11627.
(56) Gusarov, S.; Ziegler, T.; Kovalenko, A. Self-Consistent Combination of the Three-Dimensional RISM Theory of Molecular Solvation with Analytical Gradients and the Amsterdam Density Functional Package. *J. Phys. Chem. A* **2006**, *110*, 6083−6090.

DOI: 10.1021/acs.jpca.9b01623
J. Phys. Chem. A 2019, 123, 4087−4093

(57) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; et al. *AMBER*; University of California: San Francisco, 2018.
(58) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247−260.
(59) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(60) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024−10035.
(61) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminf.* **2011**, *3*, 33.
(62) Bultinck, P.; Vanholme, R.; Popelier, P. L. A.; De Proft, F.; Geerlings, P. High-Speed Calculation of AIM Charges through the Electronegativity Equalization Method. *J. Phys. Chem. A* **2004**, *108*, 10359−10366.
(63) Bultinck, P.; Langenaeker, W.; Lahorte, P.; De Proft, F.; Geerlings, P.; Van Alsenoy, C.; Tollenaere, J. P. The Electronegativity Equalization Method II: Applicability of Different Atomic Charge Schemes. *J. Phys. Chem. A* **2002**, *106*, 7895−7901.
(64) Geidl, S.; Bouchal, T.; Raček, T.; Svobodová Vařeková, R.; Hejret, V.; Křenek, A.; Abagyan, R.; Koča, J. High-Quality and Universal Empirical Atomic Charges for Chemoinformatics Applications. *J. Cheminf.* **2015**, *7*, 59.
(65) Rappe, A. K.; Goddard, W. A. Charge Equilibration for Molecular Dynamics Simulations. *J. Phys. Chem.* **1991**, *95*, 3358−3363.
(66) Chen, J.; Martínez, T. J. QTPIE: Charge Transfer with Polarization Current Equalization. A Fluctuating Charge Model with Correct Asymptotics. *Chem. Phys. Lett.* **2007**, *438*, 315−320.
(67) Imai, T. Molecular theory of partial molar volume and its applications to biomolecular systems. *Condens. Matter Phys.* **2007**, *10*, 343−361.
(68) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(69) Stewart, J. J. P. *MOPAC; Stewart Computational Chemistry*, Colorado Springs, CO, 2016.

DOI: 10.1021/acs.jpca.9b01623
J. Phys. Chem. A 2019, 123, 4087−4093